

Schauder estimates for solutions of linear parabolic integro-differential equations

Tianling Jin and Jingang Xiong

October 7, 2014

Abstract

We prove optimal pointwise Schauder estimates in the spatial variables for solutions of linear parabolic integro-differential equations. Optimal Hölder estimates in space-time for those spatial derivatives are also obtained.

1 Introduction

Integro-differential equations appear naturally when studying discontinuous stochastic process. In a series of papers of Caffarelli-Silvestre [5, 6, 7], regularities of solutions of fully nonlinear integro-differential elliptic equations such as Hölder estimates, Cordes-Nirenberg type estimates and Evans-Krylov theorem were established. Regularity for parabolic integro-differential equations has been also studied, e.g., in [8, 9, 10, 13, 14, 23] and many others. In this paper, we prove optimal pointwise Schauder estimates in the spatial variables for solutions of linear parabolic integro-differential equations. In general, we can not expect any interior continuity of the derivative of local solutions in the time variable even for the fractional heat equation $u_t + (-\Delta)^{\sigma/2}u = 0$ without extra assumptions; see example 2.4.1 in [10].

We consider the linear parabolic integro-differential equation

$$u_t(x, t) - Lu(x, t) = f(x, t) \quad \text{in } B_5 \times (-5^\sigma, 0], \quad (1.1)$$

where

$$Lu(x) := \int_{\mathbb{R}^n} \delta u(x, y; t) K(x, y; t) dy, \quad (1.2)$$

$\delta u(x, y; t) = u(x + y, t) + u(x - y, t) - 2u(x, t)$ and $K(x, y; t)$ is a positive kernel.

We will restrict our attention to symmetric kernels which satisfy

$$K(x, y; t) = K(x, -y; t). \quad (1.3)$$

This assumption is somewhat implicit in the expression (1.1). We also assume that the kernels are uniformly elliptic

$$\frac{(2-\sigma)\lambda}{|y|^{n+\sigma}} \leq K(x, y; t) \leq \frac{(2-\sigma)\Lambda}{|y|^{n+\sigma}} \quad (1.4)$$

for some $\sigma \in (0, 2)$, $0 < \lambda \leq \Lambda < \infty$, which is an essential assumption leading to local regularizations. Finally, we suppose that the kernels are C^1 away from the origin and satisfy

$$|\nabla_y K(x, y; t)| \leq \frac{\Lambda}{|y|^{n+\sigma+1}}, \quad (1.5)$$

and in certain cases we assume more that the kernels are C^2 away from the origin and satisfy

$$|\nabla_y^2 K(x, y; t)| \leq \frac{\Lambda}{|y|^{n+\sigma+2}}. \quad (1.6)$$

These smoothness assumptions are usually used to reduce the influence of the boundary data in the exterior domain, and one of the consequences is that the solutions of translation invariant (or “constant coefficients”) equations will have high regularity. Moreover, the conditions (1.5) and (1.6) are scaling invariant, which will be used in our perturbative arguments. We say that a kernel $K \in \mathcal{L}_0(\lambda, \Lambda, \sigma)$ if K satisfies (1.3) and (1.4), and $K \in \mathcal{L}_1(\lambda, \Lambda, \sigma)$ if K satisfies (1.3), (1.4) and (1.5). If in addition that K satisfies (1.6), then we say that $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$.

In this paper, all the solutions of nonlocal equations are understood in the viscosity sense, where the definitions of such solutions and their many properties can be found in [5] for elliptic equations and in [9] for parabolic equations. One may also consider *a priori* estimates for solutions of (1.1), i.e., assuming a smooth function u satisfies (1.1). To obtain pointwise Schauder estimates for solutions of (1.1) at $x = 0$, we assume that the kernel satisfies

$$\int_{\mathbb{R}^n} |K(x, y; t) - K(0, y; 0)| \min(|y|^2, r^2) dy \leq \Lambda(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) r^{2-\sigma} \quad (1.7)$$

for all $r \in (0, 1]$, $(x, t) \in B_5 \times (-5^\sigma, 0]$. (1.7) means that K is Hölder continuous at $(x, t) = (0, 0)$ in some integral sense. If $|K(x, y; t) - K(0, y; 0)| \leq \Lambda(2-\sigma)(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}})|y|^{-n-\sigma}$, then one can check that (1.7) holds. Meanwhile, we also assume that the right-hand side $f(x, t)$ is Hölder continuous at $(x, t) = (0, 0)$, i.e.,

$$|f(x, t) - f(0, 0)| \leq M_f(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) \quad \text{and} \quad |f(x, t)| \leq M_f \quad (1.8)$$

for all $(x, t) \in B_5 \times (-5^\sigma, 0]$ with some nonnegative constant M_f .

For a real number s , $[s]$ denotes the largest integer which is less than or equals to s . Our main result is the following optimal pointwise Schauder estimate in spatial variables for solutions of (1.1) with $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$.

Theorem 1.1. *Let $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$ with $2 > \sigma \geq \sigma_0 > 0$. Let $\alpha \in (0, 1)$ such that $|\sigma + \alpha - j| \geq \varepsilon_0$ for some $\varepsilon_0 > 0$, where $j = 1, 2, 3$. Suppose that (1.7) and (1.8) hold. If u is a viscosity solution of (1.1), then there exists a polynomial $P(x)$ of degree $[\sigma + \alpha]$ such that for $x \in B_1$*

$$\begin{aligned} |u(x, 0) - P(x)| &\leq C (\|u\|_{L^\infty(\mathbb{R}^n \times (-5\sigma, 0])} + M_f) |x|^{\sigma+\alpha}; \\ |\nabla^j P(0)| &\leq C (\|u\|_{L^\infty(\mathbb{R}^n \times (-5\sigma, 0])} + M_f), \quad j = 0, \dots, [\sigma + \alpha], \end{aligned} \tag{1.9}$$

where C is a positive constant depending only on $\lambda, \Lambda, n, \sigma_0, \alpha$ and ε_0 .

The constant C in (1.9) does not depend on σ , and thus, does not blow up as $\sigma \rightarrow 2$. But it blows up as $\sigma + \alpha$ approaching to integers. The condition that $\sigma + \alpha$ is not an integer is necessary even for (elliptic) fractional Laplacian equation $(-\Delta)^{\sigma/2} u = f$; see, e.g., Chapter V in [24].

Various Schauder estimates for solutions of some linear elliptic nonlocal equations were obtained before in, e.g., [1, 2, 11, 16] and global Schauder estimates for some linear parabolic nonlocal equations with non-symmetric kernels were obtained in [22] using probabilistic arguments, compared to which a feature of our estimate (1.9) in Theorem 1.1 is that the solution u of (1.1) is precisely of $C^{\sigma+\alpha}$ at $x = 0$ provided f is C^α at $x = 0$.

In the case of second order parabolic equations, if the coefficients are of C_x^α in x and only measurable in the time variable, then for a solution u of such equations, its second order spatial derivatives $\nabla_x^2 u$ are of $C_{x,t}^{\alpha, \alpha/2}$. Such results and related ones can be found in, e.g., [3, 12, 15, 17, 20, 21, 25]. Similar optimal interior Hölder estimates in space-time for spatial derivatives of solutions of (1.1) will follow from Theorem 1.1; see Corollary 2.6 and Corollary 2.7 in Section 2.3. In Theorem 1.1, we require that K and f have regularity in t at $t = 0$ as well, which is needed in our compactness arguments for weak limits of nonlocal parabolic operators.

One common difficulty in approximation arguments to obtain regularities of solutions of nonlocal equations is to control the error of the tails at infinity, which results in a slight loss of regularity compared to second order equations, especially in the case when $\sigma + \alpha > 1$ with $\sigma < 1$, and in the case $\sigma + \alpha > 2$. In this paper, we will approximate the genuine solution by solutions of “constant coefficients” equations instead of polynomials, which is inspired by [4, 19]. In this way, we do not need to take care of the tails at infinity, that leads to the optimal regularity. The only place where (1.5) or (1.6) is used is to obtain higher regularity of solutions of those corresponding “constant coefficients” equations.

In the following section, we prove the optimal pointwise Schauder estimates (1.9). We first establish high regularity for solutions of translation invariant equations in Section 2.1, which is the only place that we require K is C^2 away from the origin especially for $\sigma + \alpha \in (2, 3)$. In Section 2.2 we use perturbative arguments to prove Theorem 1.1. Section 2.3 is on the Hölder estimates in space-time for those spatial derivatives. In the Appendix, we recall some definitions and notions of nonlocal operators from [6, 10], and establish two approximation lemmas for our own purposes, which are variants of those in [6, 10].

Acknowledgements: We would like to thank Professor Luis Silvestre for many useful discussions and suggestions. We also thank Professor YanYan Li for his interests and constant encouragement.

Tianling Jin was supported in part by NSF grant DMS-1362525. Jingang Xiong was supported in part by the First Class Postdoctoral Science Foundation of China (No. 2012M520002) and Beijing Municipal Commission of Education for the Supervisor of Excellent Doctoral Dissertation (20131002701).

2 Optimal pointwise Schauder estimates in spatial variables

2.1 Translation invariant equations

In this section, we first establish good regularity on the solutions of translation invariant equations, which is similar to “constant coefficients” equation in the case of second order equations.

Proposition 2.1. *Suppose the kernel $K(y) \in \mathcal{L}_i(\lambda, \Lambda, \sigma)$ with $2 - \sigma_2 \geq \sigma \geq \sigma_1 > 1$ for some $\sigma_2 > 0$, where $i = 1$ or 2 . If v is a viscosity solution of*

$$v_t(x, t) - \int_{\mathbb{R}^n} \delta v(x, y; t) K(y) dy = g(x, t) \quad \text{in } B_8 \times (-8^\sigma, 0],$$

where $g(\cdot, t) \in C_x^i(B_8)$ for all $t \in (-8^\sigma, 0]$, then there exists a positive constant c_1 depending only on $n, \lambda, \Lambda, \sigma_1$ such that

$$\sup_{t \in (-1, 0)} \|v(\cdot, t)\|_{C^{1+i}(B_1)} \leq c_1 (\|v\|_{L^\infty(\mathbb{R}^n \times (-8^\sigma, 0])} + \sup_{t \in (-8^\sigma, 0)} \|g(\cdot, t)\|_{C_x^i(B_8)}); \quad (2.1)$$

and there exists another positive constant c_2 depending only on $n, \lambda, \Lambda, \sigma_1, \sigma_2$ such that

$$\sup_{t \in (-1, 0)} \|v(\cdot, t)\|_{C^{\sigma+i}(B_1)} \leq c_2 (\|v\|_{L^\infty(\mathbb{R}^n \times (-8^\sigma, 0])} + \sup_{t \in (-8^\sigma, 0)} \|g(\cdot, t)\|_{C_x^i(B_8)}). \quad (2.2)$$

This proposition will follow from the next lemma and standard integration by part techniques.

Lemma 2.2. *Let the kernel $K(y) \in \mathcal{L}_0(\lambda, \Lambda, \sigma)$ with $\sigma \geq \sigma_1 > 1$. Suppose that there exist two positive constants \tilde{c} , and $\bar{\alpha} \leq 1$ satisfying $|\bar{\alpha} - \sigma + 1| \geq \varepsilon_0$ for some $\varepsilon_0 > 0$, such that for every viscosity solution u of*

$$u_t(x, t) - \int_{\mathbb{R}^n} \delta u(x, y; t) K(y) dy = 0 \quad \text{in } B_5 \times (-5^\sigma, 0],$$

there holds

$$\|\nabla_x u(\cdot, 0)\|_{C^{0, \bar{\alpha}}(B_1)} \leq \tilde{c} \|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])}. \quad (2.3)$$

Then there exists a positive constant C depending only on $n, \lambda, \Lambda, \bar{\alpha}, \tilde{c}, \varepsilon_0$ and σ_1 such that for every viscosity solution v of

$$v_t(x, t) - \int_{\mathbb{R}^n} \delta v(x, y; t) K(y) dy = h(x, t) \quad \text{in } B_5 \times (-5^\sigma, 0],$$

there holds

$$\|\nabla_x v(\cdot, 0)\|_{C^\beta(B_1)} \leq C \left(\|v\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + \|h\|_{L^\infty(B_5 \times (-5^\sigma, 0])} \right), \quad (2.4)$$

where $\beta = \min(\sigma - 1, \bar{\alpha})$.

Note that it follows from [23] that our assumption (2.3) indeed holds for some $\bar{\alpha} > 0$. If we assume $K(y) \in \mathcal{L}_1(\lambda, \Lambda, \sigma)$ with $\sigma > 1$, then by using Theorem 6.2 in [9], integration by part techniques and Lemma 2.2 itself, we will see in the proof of Proposition 2.1 that (2.3) actually holds with $\bar{\alpha} = 1$, and thus, (2.4) holds with $\beta = \sigma - 1$.

Proof. We can assume that $\|h(x, t)\|_{L^\infty(B_5 \times (-5^\sigma, 0])} + \|v\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} \leq 1$. Let $\rho = 1/2$. For $k = 0, 1, 2, \dots$, let $Q_k = B_{\rho^k} \times (-\rho^{k\sigma}, 0]$ and v_k be the solution of the following translation invariant equation

$$\begin{aligned} \partial_t v_k(x, t) - \int_{\mathbb{R}^n} \delta v_k(x, y; t) K(y) dy &= 0 \quad \text{in } Q_k, \\ v_k &= v \quad \text{in } ((\mathbb{R}^n \setminus B_{\rho^k}) \times [-\rho^{k\sigma}, 0]) \cup (\mathbb{R}^n \times \{t = -\rho^{k\sigma}\}). \end{aligned}$$

The existence and uniqueness of such v_k is guaranteed by Theorem 3.3 in [9]. Then we have by the maximum principle,

$$\|v_k - v\|_{L^\infty(\mathbb{R}^n \times [-\rho^{k\sigma}, 0])} \leq \rho^{\sigma k}$$

and thus, by the maximum principle again,

$$\|v_k - v_{k+1}\|_{L^\infty(\mathbb{R}^n \times [-\rho^{(k+1)\sigma}, 0])} \leq \|v_k - v\|_{L^\infty(\mathbb{R}^n \times [-\rho^{k\sigma}, 0])} \leq \rho^{\sigma k}.$$

Let $w_{k+1} = v_{k+1} - v_k$. It follows from the assumption estimate (2.3) that for $x \in B_{\rho^{k+2}}$,

$$\begin{aligned} |\nabla_x w_{k+1}(x, 0)| &\leq C \rho^{(\sigma-1)k} \\ |w_{k+1}(x, 0) - w_{k+1}(0, 0) - \nabla_x w_{k+1}(0, 0) \cdot x| &\leq C \rho^{(\sigma-1-\bar{\alpha})k} |x|^{1+\bar{\alpha}}. \end{aligned}$$

Thus, for $\rho^{i+2} \leq |x| < \rho^{i+1}$, if we let $w = v - v_0$, then

$$\begin{aligned} &|w(x, 0) - \sum_{l=1}^{\infty} w_l(0, 0) - \sum_{l=1}^{\infty} \nabla_x w_l(0, 0) \cdot x| \\ &\leq |w(x, 0) - \sum_{l=1}^i w_l(x, 0)| + |\sum_{l=1}^i w_l(x, 0) - \sum_{l=1}^i w_l(0, 0) - \sum_{l=1}^i \nabla_x w_l(0, 0) \cdot x| \\ &\quad + |\sum_{l=i+1}^{\infty} w_l(0, 0)| + |\sum_{l=i+1}^{\infty} \nabla_x w_l(0, 0) \cdot x| \\ &\leq \rho^{\sigma i} + C|x|^{1+\bar{\alpha}} \sum_{l=1}^i \rho^{(\sigma-1-\bar{\alpha})l} + C \sum_{l=i+1}^{\infty} \rho^{\sigma l} + C|x| \sum_{l=i+1}^{\infty} \rho^{(\sigma-1)l} \\ &\leq C|x|^{\beta+1}, \end{aligned} \quad (2.5)$$

where $\beta = \min(\sigma - 1, \bar{\alpha})$, and C depends only on $n, \lambda, \Lambda, \bar{\alpha}, \tilde{c}, \varepsilon_0$ and σ_1 . Meanwhile, it follows from the assumption estimate (2.3) that

$$|v_0(x, 0) - v_0(0, 0) - \nabla_x v_0(0, 0)x| \leq C|x|^{1+\bar{\alpha}}.$$

This finishes the proof. \square

Proof of Proposition 2.1. First of all, we know from Theorem 6.2 in [9] that $\nabla_x v$ is local Hölder continuous in space-time. We will use integration by parts techniques which can be found in [5]. Let η_1 be a smooth cut-off function supported in B_7 and $\eta_1 \equiv 1$ in B_6 . Let $w_1 = \nabla_x(\eta_1 v)$. Then it satisfies in viscosity sense that

$$\begin{aligned} \partial_t w_1(x, t) - \int_{\mathbb{R}^n} \delta w_1(x, y; t) K(y) dy \\ = - \int_{\mathbb{R}^n} ((1 - \eta_1)v)(x + y, t) \nabla_y K(y) dy + \nabla_x g(x, t) \quad \text{in } B_5 \times (-5^\sigma, 0]. \end{aligned}$$

Thus, if $K(y) \in \mathcal{L}_1(\lambda, \Lambda, \sigma)$, it follows from Lemma 2.2 that w_1 is $C^{1+\beta}$ in x for some $\beta > 0$. Thus, we have C^2 estimate in x (2.1) for v . This implies that the assumption estimate (2.3) in Lemma 2.2 is satisfied with $\bar{\alpha} = 1$ if $K(y) \in \mathcal{L}_1(\lambda, \Lambda, \sigma)$. Now, we apply Lemma 2.2 once more to the equation of w_1 . If we choose $\bar{\alpha} = 1$ we have that w_1 is C^σ in x , from which (2.2) follows. This proves the case of $i = 1$.

If $K(y) \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$, then we take another smooth cut-off function η_2 supported in B_4 and $\eta_2 \equiv 1$ in B_3 . Let $w_2 = \nabla_x(\eta_2 w_1)$. Then it satisfies

$$\begin{aligned} \partial_t w_2(x, t) - \int_{\mathbb{R}^n} \delta w_2(x, y; t) K(y) dy = - \int_{\mathbb{R}^n} ((1 - \eta_2)w_1)(x + y) \nabla_y K(y) dy \\ + \int_{\mathbb{R}^n} ((1 - \eta_1)v)(x + y) \nabla_y^2 K(y) dy + \nabla_x^2 g(x, t) \quad \text{in } B_2 \times (-2^\sigma, 0]. \end{aligned}$$

Thus, (2.1) and (2.2) follow as before. This proves the case of $i = 2$. \square

Similarly, for $0 < \sigma_0 \leq \sigma \leq 1$, we have

Lemma 2.3. *Let the kernel $K(y) \in \mathcal{L}_0(\lambda, \Lambda, \sigma)$ with $0 < \sigma_0 \leq \sigma \leq 1$. Suppose that there exist two positive constants \tilde{c} , and $\bar{\alpha} \leq 1$ satisfying $|\bar{\alpha} - \sigma| \geq \varepsilon_0$ for some $\varepsilon_0 > 0$ such that for every viscosity solution u of*

$$u_t(x, t) - \int_{\mathbb{R}^n} \delta u(x, y; t) K(y) dy = 0 \quad \text{in } B_5 \times (-5^\sigma, 0],$$

there holds

$$\|u(\cdot, 0)\|_{C^{0, \bar{\alpha}}(B_1)} \leq \tilde{c} \|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])}. \quad (2.6)$$

Then there exists a positive constant C depending only on $n, \lambda, \Lambda, \bar{\alpha}, \tilde{c}, \varepsilon_0$ and σ_0 such that for every viscosity solution v of

$$v_t(x, t) - \int_{\mathbb{R}^n} \delta v(x, y; t) K(y) dy = h(x, t) \quad \text{in } B_5 \times (-5^\sigma, 0],$$

there holds

$$\|v(\cdot, 0)\|_{C^\beta(B_1)} \leq C \left(\|v\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + \|h\|_{L^\infty(B_5 \times (-5^\sigma, 0])} \right), \quad (2.7)$$

where $\beta = \min(\sigma, \bar{\alpha})$.

Note that it follows from Theorem 6.1 in [9] that our assumption (2.6) indeed holds for some $\bar{\alpha} > 0$. If we assume $K(y) \in \mathcal{L}_1(\lambda, \Lambda, \sigma)$, then it follows from Theorem 6.2 in [9] that (2.6) actually holds with $\bar{\alpha} = 1$.

Proposition 2.4. *Suppose the kernel $K(y) \in \mathcal{L}_i(\lambda, \Lambda, \sigma)$ with $1 \geq \sigma > \sigma_0 > 0$, where $i = 1$ or 2 . If v is a viscosity solution of*

$$v_t(x, t) - \int_{\mathbb{R}^n} \delta v(x, y; t) K(y) dy = g(x, t) \quad \text{in } B_8 \times (-8^\sigma, 0],$$

where $g(\cdot, t) \in C_x^i(B_8)$ for all $t \in [-8^\sigma, 0]$, then there exist a positive constant c_1 depending only on $n, \lambda, \Lambda, \sigma_0$ such that

$$\sup_{t \in (-1, 0)} \|v(\cdot, t)\|_{C^i(B_1)} \leq c_1 (\|v\|_{L^\infty(\mathbb{R}^n \times (-8^\sigma, 0])} + \sup_{t \in (-8^\sigma, 0)} \|g(\cdot, t)\|_{C_x^i(B_8)}).$$

When $\sigma \leq 1 - \varepsilon_0$ for some $\varepsilon_0 > 0$, there exist a positive constant c_2 depending only on $n, \lambda, \Lambda, \sigma_0, \varepsilon_0$ such that

$$\sup_{t \in (-1, 0)} \|v(\cdot, t)\|_{C^{\sigma+i}(B_1)} \leq c_2 (\|v\|_{L^\infty(\mathbb{R}^n \times (-8^\sigma, 0])} + \sup_{t \in (-8^\sigma, 0)} \|g(\cdot, t)\|_{C_x^i(B_8)}).$$

When $\sigma = 1$, then for all $\beta \in (0, 1)$ there exist a positive constant c_3 depending only on $n, \lambda, \Lambda, \sigma_0, \beta$ such that

$$\sup_{t \in (-1, 0)} \|v(\cdot, t)\|_{C^{\beta+i}(B_1)} \leq c_3 (\|v\|_{L^\infty(\mathbb{R}^n \times (-8^\sigma, 0])} + \sup_{t \in (-8^\sigma, 0)} \|g(\cdot, t)\|_{C_x^i(B_8)}).$$

The proofs of Lemma 2.3 and Proposition 2.4 are very similar to those of Lemma 2.2 and Proposition 2.1, respectively, and we leave them to the readers.

2.2 Proof of the main theorem

Now we are in position to prove Theorem 1.1 by approximations.

Proof of Theorem 1.1. The strategy of the proof is to find a sequence of approximation solutions which are sufficiently regular, and the error between the genuine solution and the approximation solutions can be controlled in a desired rate.

We may assume that $\|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + M_f \leq 1$. We claim that we can inductively find a sequence of functions w_i , $i = 0, 1, 2, \dots$, such that for all i ,

$$\partial_t \sum_{l=0}^i w_l(x, t) - L_0 \sum_{l=0}^i w_l(x, t) = f(0, 0) \quad \text{in } B_{4 \cdot 5^{-i}} \times (-4^\sigma \cdot 5^{-i\sigma}, 0], \quad (2.8)$$

and

$$(u - \sum_{l=0}^i w_l)(5^{-i}x, 5^{-i\sigma}t) = 0 \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\sigma, 0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}), \quad (2.9)$$

and

$$\|u - \sum_{l=0}^i w_l\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma 5^{-i\sigma}, 0])} \leq 5^{-(\sigma+\alpha)(i+1)}, \quad (2.10)$$

and

$$\begin{aligned} \|w_i\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma 5^{-i\sigma}, 0])} &\leq 5^{-(\sigma+\alpha)i} \\ \|\nabla_x w_i\|_{L^\infty(B_{(4-\tau) \cdot 5^{-i}} \times [(-4^\sigma + \tau^\sigma) 5^{-i\sigma}, 0])} &\leq c_2 5^{-(\sigma+\alpha-1)i} \tau^{-1} \\ \|\nabla_x^2 w_i\|_{L^\infty(B_{(4-\tau) \cdot 5^{-i}} \times [(-4^\sigma + \tau^\sigma) 5^{-i\sigma}, 0])} &\leq c_2 5^{-(\sigma+\alpha-2)i} \tau^{-2}, \\ \|\nabla_x^3 w_i\|_{L^\infty(B_{(4-\tau) \cdot 5^{-i}} \times [(-4^\sigma + \tau^\sigma) 5^{-i\sigma}, 0])} &\leq c_2 5^{-(\sigma+\alpha-3)i} \tau^{-3} \quad (\text{if } \sigma > 2 - \alpha), \end{aligned} \quad (2.11)$$

and

$$[u - \sum_{l=0}^i w_l]_{C^{\alpha_1}(B_{(4-3\tau) \cdot 5^{-i}} \times [(-4^\sigma + 2\tau^\sigma) 5^{-i\sigma}, 0])} \leq 8c_1 5^{i\alpha_1 - (\sigma+\alpha)(i+1)} \tau^{-4}, \quad (2.12)$$

where τ is an arbitrary constant in $(0, 1)$, $c_1 > 0$ and $\alpha_1 \in (0, 1)$ are positive constants depending only on $\lambda, \Lambda, n, \sigma_0$, and $c_2 > 0$ additionally depends on α . Then, Theorem 1.1 follows from this claim and standard arguments. Indeed, as in (2.5), we have, for $5^{-(i+1)} \leq |x| < 5^{-i}$,

$$\begin{aligned} |u(x, 0) - \sum_{l=0}^{\infty} w_l(0, 0)| &\leq C_1 |x|^{\sigma+\alpha} \quad \text{when } \sigma + \alpha < 1, \\ |u(x, 0) - \sum_{l=0}^{\infty} w_l(0, 0) - \sum_{l=0}^{\infty} \nabla_x w_l(0, 0) \cdot x| &\leq C_2 |x|^{\sigma+\alpha} \quad \text{when } 1 < \sigma + \alpha < 2. \end{aligned}$$

When $2 < \sigma + \alpha < 3$, we have, for $5^{-(i+1)} \leq |x| < 5^{-i}$,

$$\begin{aligned}
& |u(x, 0) - \sum_{l=0}^{\infty} w_l(0, 0) - \sum_{l=0}^{\infty} \nabla_x w_l(0, 0) \cdot x - \sum_{l=0}^{\infty} \frac{1}{2} x^T \nabla_x^2 w_l(0, 0) x| \\
& \leq |u(x, 0) - \sum_{l=0}^i w_l(x, 0)| \\
& \quad + |\sum_{l=0}^i w_l(x, 0) - \sum_{l=0}^i w_l(0, 0) - \sum_{l=0}^i \nabla_x w_l(0, 0) \cdot x - \sum_{l=0}^i \frac{1}{2} x^T \nabla_x^2 w_l(0, 0) x| \\
& \quad + |\sum_{l=i+1}^{\infty} w_l(0, 0)| + |\sum_{l=i+1}^{\infty} \nabla_x w_l(0, 0) \cdot x| + \frac{1}{2} |\sum_{l=i+1}^{\infty} x^T \nabla_x^2 w_l(0, 0) x| \\
& \leq 5^{-(\sigma+\alpha)(i+1)} + 2c_2 |x|^3 \sum_{l=0}^i 5^{-(\sigma+\alpha-3)l} + \sum_{l=i+1}^{\infty} 5^{-(\sigma+\alpha)l} + |x| \sum_{l=i+1}^{\infty} c_2 5^{-(\sigma+\alpha-1)l} \\
& \quad + |x|^2 \sum_{l=i+1}^{\infty} c_2 5^{-(\sigma+\alpha-2)l} \\
& \leq C_3 |x|^{\sigma+\alpha}.
\end{aligned}$$

Note that we used $|\sigma + \alpha - j| \geq \varepsilon_0$ for $j=1,2,3$ in obtaining C_1, C_2, C_3 , which actually blow up at a rate of $O(|\sigma + \alpha - j|^{-1})$ as $\sigma + \alpha \rightarrow j \in \{1, 2, 3\}$. The estimate (1.9) is proved using the claim.

Now we are left to prove this claim. Before we provide the detailed proof, we would like to first mention the idea and the structure of (2.8)-(2.12):

- Solving (2.8) and (2.9) inductively is how we construct this sequence of functions $\{w_i\}$.
- (2.10) will follow from the approximation lemmas in the appendix, where (2.12) will be used.
- (2.11) will follow from (2.10), maximum principles and the estimates in Proposition 2.1 and Proposition 2.4.

The proof of the above claim is by induction, and it consists of three steps.

Step 1: Normalization and rescaling.

Let w_0 be the viscosity solution of

$$\begin{aligned}
\partial_t w_0 - L_0 w_0 &= f(0, 0) \quad \text{in } B_4 \times (-4^\sigma, 0] \\
w_0 &= u \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\sigma, 0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}),
\end{aligned} \tag{2.13}$$

where

$$L_0 w = \int_{\mathbb{R}^n} \delta w(x, y; t) K(0, y; 0) dy.$$

We also think of $w_0 \equiv u$ in $\mathbb{R}^n \times (-5^\sigma, -4^\sigma)$. Then by comparison principles,

$$\|w_0\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, 0])} \leq c_0(\|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + \|f\|_{L^\infty(B_5 \times (-5^\sigma, 0])}), \quad (2.14)$$

where c_0 is a positive constant depending only on $n, \lambda, \Lambda, \sigma_0$. By normalization, we may assume that

$$\|w_0\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, 0])} \leq 1, \quad \|u\|_{L^\infty(\mathbb{R}^n \times [-5^\sigma, 0])} + \|f\|_{L^\infty(B_5 \times [-5^\sigma, 0])} \leq 1.$$

For some universal small positive constant $\gamma < 1$, which will be chosen in (2.23), we also may assume that $|f(x, t) - f(0, 0)| \leq \gamma(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}})$ in $B_5 \times (-5^\sigma, 0]$ and

$$\int_{\mathbb{R}^n} |K(x, y; t) - K(0, y; 0)| \min(|y|^2, r^2) dy \leq \gamma(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) r^{2-\sigma} \quad (2.15)$$

for all $r \in (0, 1]$, $(x, t) \in B_5 \times (-5^\sigma, 0]$. This can be achieved by the scaling for $r < 1$ small that if we let

$$\begin{aligned} \tilde{K}(x, y; t) &= r^{n+\sigma} K(rx, ry; r^\sigma t) \in \mathcal{L}_2(\lambda, \Lambda, \sigma), \\ \tilde{u}(x, t) &= u(rx, r^\sigma t), \\ \tilde{f}(x, t) &= r^\sigma f(rx, r^\sigma t), \end{aligned}$$

then we see that

$$\tilde{u}_t(x, t) - \tilde{L}\tilde{u}(x, t) = \tilde{f}(x, t) \quad \text{in } B_5 \times (-5^\sigma, 0]$$

where

$$\tilde{L}\tilde{u}(x, t) := \int_{\mathbb{R}^n} \delta \tilde{u}(x, y; t) \tilde{K}(x, y, t) dy.$$

Thus

$$|\tilde{f}(x, t) - \tilde{f}(0, 0)| \leq M_f r^{\sigma+\alpha} (|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) \leq \gamma(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) \leq 10\gamma \quad \text{for small } r,$$

in $B_5 \times (-5^\sigma, 0]$ and

$$\begin{aligned} \int_{\mathbb{R}^n} |\tilde{K}(x, y; t) - \tilde{K}(0, y; 0)| \min(|y|^2, s^2) dy &\leq 2\Lambda r^\alpha (|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) s^{2-\sigma} \\ &\leq \gamma(|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) s^{2-\sigma} \end{aligned}$$

for all $s \in (0, 1]$, $(x, t) \in B_5 \times (-5^\sigma, 0]$. It follows that $(\|\cdot\|_*)$ is defined in (A.1) in the Appendix

$$\|\tilde{L} - \tilde{L}_0\|_* \leq 50\gamma \quad \text{in } B_5 \times (-5^\sigma, 0].$$

Indeed, for $(x, t) \in B_5 \times (-5^\sigma, 0]$, $\|h\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} \leq M$ and $|h(x + y, t) - h(x, t) - y \cdot \nabla_x h(x, t)| \leq M|y|^2$ for every $y \in B_1$, we have

$$\begin{aligned}
& \|\tilde{L} - \tilde{L}_0\|_* \\
& \leq \sup_{(x,t),h} \frac{1}{1+M} \int_{\mathbb{R}^n} |\delta h(x, y; t)| |\tilde{K}(x, y; t) - \tilde{K}(0, y; 0)| dy \\
& \leq \sup_{(x,t)} \frac{M}{1+M} \left(\int_{B_1} |y|^2 |\tilde{K}(x, y; t) - \tilde{K}(0, y; 0)| dy \right. \\
& \quad \left. + 4 \int_{\mathbb{R}^n \setminus B_1} |\tilde{K}(x, y; t) - \tilde{K}(0, y; 0)| dy \right) \\
& < 50\gamma.
\end{aligned} \tag{2.16}$$

Step 2: Prove the claim for $i = 0$.

Let w_0 be the one in Step 1. It follows from Proposition 2.1 and Proposition 2.4 that there exists a positive constant c_2 depending only on $\lambda, \Lambda, n, \sigma_0, \alpha$ such that

$$\begin{aligned}
& \|w_0\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, 0])} \leq 1 \\
& \|\nabla_x w_0\|_{L^\infty(B_{4-\tau} \times [-4^\sigma + \tau^\sigma, 0])} \leq c_2 \tau^{-1} \\
& \|\nabla_x^2 w_0\|_{L^\infty(B_{4-\tau} \times [-4^\sigma + \tau^\sigma, 0])} \leq c_2 \tau^{-2} \\
& \|\nabla_x^3 w_0\|_{L^\infty(B_{4-\tau} \times [-4^\sigma + \tau^\sigma, 0])} \leq c_2 \tau^{-3} \text{ (if } \sigma > 2 - \alpha \text{)}.
\end{aligned} \tag{2.17}$$

For $\tau \in (0, 1]$, it follows from Theorem 6.1 in [9] (see [5] for the elliptic case), standard scaling and covering (contributing at most a factor of $4/\tau$) argument that there exist constants $\alpha_1 \in (0, 1)$, $c_1 > 0$, depending only on $n, \lambda, \Lambda, \sigma_0$, such that

$$\|u\|_{C^{\alpha_1}(B_{4-\tau} \times [-4^\sigma + \tau^\sigma, 0])} \leq c_1 \tau^{-\alpha_1 - 1}. \tag{2.18}$$

Let us set up to apply the first approximation lemma in the Appendix, Lemma A.1. Let $\varepsilon = 5^{-(\sigma+\alpha)}$ and $M_1 = 1$ and let us fixed a modulus continuity $\rho(s) = s^{\alpha_1}$. Then for these ρ, ε, M , there exist η_1 (small) and R (large) so that Lemma A.1 holds. We can rescale the equation of u so that it holds in a very large cylinder containing $B_{2R} \times [-(2R)^\sigma, 0]$ and $|u(x, t) - u(y, s)| \leq \rho(|x - y| \vee |t - s|)$ for every $(x, t) \in (B_R \setminus B_4) \times [-4^\sigma, 0]$ and $(y, s) \in (\mathbb{R}^n \setminus B_4) \times [-4^\sigma, 0] \cup \mathbb{R}^n \times \{s = -4^\sigma\}$. The latter one can be done due to (2.18). And we will choose $\gamma < \eta_1/50$ in (2.23). Then we can conclude from Lemma A.1 that

$$\|u - w_0\|_{L^\infty(B_4 \times [-4^\sigma, 0])} \leq \varepsilon = 5^{-(\sigma+\alpha)},$$

and thus,

$$\|u - w_0\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, 0])} \leq \|u - w_0\|_{L^\infty(B_4 \times [-4^\sigma, 0])} \leq \varepsilon = 5^{-(\sigma+\alpha)}.$$

This proves (2.8), (2.9), (2.10) and (2.11) hold for $i = 0$.

Moreover,

$$|(u - w_0)_t - L(u - w_0)| \leq 10\gamma + (c_2 + 4)\gamma\tau^{-\sigma} \quad (2.19)$$

in $B_{4-2\tau} \times (-4^\sigma + \tau^\sigma, 0]$ in viscosity sense. Indeed, let $t_0 \in (0, 1)$ and we smooth w_0 by using a mollifier $\eta_\varepsilon(x, t)$, and let $g_\varepsilon = \eta_\varepsilon * w_0$ (thinking of $w_0 \equiv u$ in $\mathbb{R}^n \times [-5^\sigma, -4^\sigma]$). Let w_0^ε be the solution of

$$\begin{aligned} \partial_t w_0^\varepsilon - L_0 w_0^\varepsilon &= f(0, 0) \quad \text{in } B_4 \times (-4^\sigma, -t_0] \\ w_0^\varepsilon &= g_\varepsilon \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\sigma, -t_0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}). \end{aligned}$$

It follows from Theorem 4.1 in [10] that $\partial_t v_\varepsilon$ is Hölder continuous in space-time. Thus,

$$(u - w_0^\varepsilon)_t - L(u - w_0^\varepsilon) = f(x, t) - f(0, 0) + \int_{\mathbb{R}^n} \delta w_0^\varepsilon(x, y; t) (K(x, y; t) - K(0, y; 0)) dy.$$

For $(x, t) \in B_{4-2\tau} \times [-4^\sigma + \tau^\sigma, t_0]$, it follows from Proposition 2.1 and Proposition 2.4 that

$$\begin{aligned} & \int_{\mathbb{R}^n} |\delta w_0^\varepsilon(x, y; t)| |K(0, y; 0) - K(x, y; t)| dy \\ & \leq \int_{B_\tau} c_2 \tau^{-2} |y|^2 |K(0, y; 0) - K(x, y; t)| dy \\ & \quad + 4 \int_{\mathbb{R}^n \setminus B_\tau} |K(0, y; 0) - K(x, y; t)| dy \\ & \leq (c_2 + 4)\gamma\tau^{-\sigma} (|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) \leq 10(c_2 + 4)\gamma\tau^{-\sigma}. \end{aligned} \quad (2.20)$$

Meanwhile, by the Hölder interior estimates, we have that w_0^ε locally uniformly converges to some continuous function w . By the stability result Theorem 5.3 in [10], w is a viscosity solution of

$$\begin{aligned} \partial_t w - L_0 w &= f(0, 0) \quad \text{in } B_4 \times (-4^\sigma, -t_0] \\ w &= w_0 \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\sigma, -t_0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}). \end{aligned}$$

Hence $w \equiv w_0$. Thus, by sending $\varepsilon \rightarrow 0$, and $t_0 \rightarrow 0$ with a standard perturbation argument (using ν/t for small ν), (2.19) holds in $B_{4-2\tau} \times (-4^\sigma + \tau^\sigma, 0]$ in viscosity sense. By the choice of γ in (2.23),

$$|(u - w_0)_t(x, t) - L(u - w_0)(x, t)| \leq 10\gamma + 10(c_2 + 4)\gamma\tau^{-\sigma} \leq 5^{-(\sigma+\alpha)}\tau^{-\sigma}.$$

It follows from the Hölder estimates (2.18) proved in [9], standard rescaling and covering arguments (contributing at most a factor of $4/\tau$) that

$$[u - w_0]_{C^{\alpha_1}(B_{4-3\tau} \times [-4^\sigma + 2\tau^\sigma, 0])} \leq \frac{4}{\tau} c_1 \tau^{-\alpha_1} (5^{-(\sigma+\alpha)}\tau^{-\sigma} + 5^{-(\sigma+\alpha)}) \leq 8c_1 5^{-(\sigma+\alpha)}\tau^{-4}.$$

This finishes the proof of (2.12) for $i = 0$.

Step 3: We assume all of (2.8), (2.9), (2.10), (2.11), and (2.12) hold up to $i \geq 0$. We will show that they all hold for $i + 1$ as well.

Let

$$W(x, t) = 5^{(\sigma+\alpha)(i+1)} \left(u - \sum_{l=0}^i w_l \right) (5^{-(i+1)}x, 5^{-(i+1)\sigma}t)$$

and

$$K^{(i+1)}(x, y; t) = 5^{-(n+\sigma)(i+1)} K(5^{-(i+1)}x, 5^{-(i+1)}y; 5^{-(i+1)\sigma}t).$$

Thus, by (2.8), we have as before

$$|W_t(x, t) - \int_{\mathbb{R}^n} \delta W(x, y; t) K^{(i+1)}(x, y; t) dy| \leq A$$

in viscosity sense in $B_{(4-2\tau)\cdot 5} \times [(-4^\sigma + \tau^\sigma)5^\sigma, 0]$, where A is a constant such that

$$\begin{aligned} A &\leq |5^{\alpha(i+1)}(f(5^{-(i+1)}x, 5^{-(i+1)\sigma}t) - f(0, 0))| \\ &+ \sum_{l=0}^i \int_{\mathbb{R}^n} 5^{(\sigma+\alpha)(i+1)} |\delta w_l(5^{-(i+1)}x, 5^{-(i+1)}y; 5^{-(i+1)\sigma}t)| \cdot \\ &\quad |K^{(i+1)}(x, y; t) - K^{(i+1)}(0, y; 0)| dy. \end{aligned}$$

Then for $(x, t) \in B_{20} \times (-20^\sigma, 0)$,

$$|5^{\alpha(i+1)}(f(5^{-(i+1)}x, 5^{-(i+1)\sigma}t) - f(0, 0))| \leq 40 \cdot \gamma,$$

and for $l = 0, 1, \dots, i$ and for $(x, t) \in B_{(4-2\tau)\cdot 5} \times [(-4^\sigma + \tau^\sigma)5^\sigma, 0]$, we have, similar to (2.20),

$$\begin{aligned} &5^{(\sigma+\alpha)(i+1)} \int_{\mathbb{R}^n} |\delta w_l(5^{-(i+1)}x, 5^{-(i+1)}y; 5^{-(i+1)\sigma}t)| \cdot \\ &\quad |K^{(i+1)}(x, y; t) - K^{(i+1)}(0, y; 0)| dy \\ &= 5^{\alpha(i+1)} \int_{\mathbb{R}^n} |\delta w_l(5^{-(i+1)}x, y; 5^{-(i+1)\sigma}t)| |K(5^{-(i+1)}x, y; 5^{-(i+1)\sigma}t) - K(0, y; 0)| dy \\ &\leq 5^{\alpha(i+1)} \int_{B_{5^{-l}\tau}} c_2 5^{-(\sigma+\alpha-2)l} \tau^{-2} |K(5^{-(i+1)}x, y; 5^{-(i+1)\sigma}t) - K(0, y; 0)| dy \\ &\quad + 5^{\alpha(i+1)} \int_{\mathbb{R}^n \setminus B_{5^{-l}\tau}} 4 \cdot 5^{-(\sigma+\alpha)l} |K(5^{-(i+1)}x, y; 5^{-(i+1)\sigma}t) - K(0, y; 0)| dy \\ &\leq \gamma c_2 \tau^{-\sigma} 5^{-\alpha l} (|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) + \gamma 4 \tau^{-\sigma} 5^{-\alpha l} (|x|^\alpha + |t|^{\frac{\alpha}{\sigma}}) \\ &\leq \gamma 40(c_2 + 4) \tau^{-\sigma} 5^{-\alpha l}. \end{aligned} \tag{2.21}$$

Thus, for $(x, t) \in B_{(4-2\tau) \cdot 5} \times [(-4^\sigma + \tau^\sigma)5^\sigma, 0]$, we have

$$\begin{aligned}
& |W_t(x, t) - \int_{\mathbb{R}^n} \delta W(x, y; t) K^{(i+1)}(x, y; t) dy| \\
& \leq 40\gamma + \gamma 40(c_2 + 4) \tau^{-\sigma} \sum_{l=0}^{\infty} 5^{-\alpha l} \\
& = \tau^{-\sigma} \left(40 + 40(c_2 + 4) \sum_{l=0}^{\infty} 5^{-\alpha l} \right) \gamma.
\end{aligned} \tag{2.22}$$

Let τ_0 be such that $-4^\sigma + 2\tau_0^\sigma < -2^\sigma$ which depends only on σ_0 . Let $\eta_2 < 5^{-(\sigma+\alpha)}$ be as in Lemma A.2 with $M_2 = 1, M_3 = 8c_1, \beta = \alpha_1$ and $\varepsilon = 5^{-(\sigma+\alpha)}$. We choose γ such that

$$\gamma < \eta_1/50 \quad \text{and} \quad \tau_0^{-2} \left(40 + 40(c_2 + 4) \sum_{l=0}^{\infty} 5^{-\alpha l} \right) \gamma \leq \eta_2. \tag{2.23}$$

By our induction hypothesis (2.10), (2.12) and (2.9),

$$\begin{aligned}
& \|W\|_{L^\infty(\mathbb{R}^n \times [-20^\sigma, 0])} \leq 1, \\
& [W]_{C^{\alpha_1}(B_{(4-3\tau) \cdot 5} \times [(-4^\sigma + 2\tau^\sigma)5^\sigma, 0])} \leq 8c_1 5^{-\alpha_1} \tau^{-4} \leq 8c_1 \tau^{-4}. \\
& W(x) = 0 \quad \text{for all } (x, t) \in (\mathbb{R}^n \setminus B_{20}) \times [-20^\tau, 0]
\end{aligned}$$

It follows from (2.22), (2.23) and the choice of τ_0 that

$$|W_t(x, t) - \int_{\mathbb{R}^n} \delta W(x, y; t) K^{(i+1)}(x, y; t) dy| \leq \eta_2 \quad \text{for } (x, t) \in B_{10} \times (-10^\sigma, 0].$$

Let v_{i+1} be the solution of

$$\begin{aligned}
& \partial_t v_{i+1} - \int_{\mathbb{R}^n} \delta v_{i+1}(x, y; t) K^{(i+1)}(0, y; 0) dy = 0 \quad \text{in } B_4 \times (-4^\sigma, 0], \\
& v_{i+1} = W \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\tau, 0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}).
\end{aligned}$$

Together with the calculation in (2.16), it follows from Lemma A.2 that

$$\|W - v_{i+1}\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, 0])} = \|W - v_{i+1}\|_{L^\infty(B_4 \times [-4^\sigma, 0])} \leq 5^{-(\sigma+\alpha)}. \tag{2.24}$$

Moreover, it follows from (2.10) that

$$\|v_{i+1}\|_{L^\infty(\mathbb{R}^n \times [-20^\sigma, 0])} \leq \|W\|_{L^\infty(\mathbb{R}^n \times [-20^\sigma, 0])} \leq 1,$$

and thus by Proposition 2.1 and Proposition 2.4 (see also (2.17))

$$\begin{aligned}\|\nabla_x v_{i+1}\|_{L^\infty(B_{4-\tau} \times [-4\sigma + \tau^\sigma, 0])} &\leq c_2 \tau^{-1}, \\ \|\nabla_x^2 v_{i+1}\|_{L^\infty(B_{4-\tau} \times [-4\sigma + \tau^\sigma, 0])} &\leq c_2 \tau^{-2}, \\ \|\nabla_x^3 v_{i+1}\|_{L^\infty(B_{4-\tau} \times [-4\sigma + \tau^\sigma, 0])} &\leq c_2 \tau^{-3} \text{ (if } \sigma > 1\text{)}.\end{aligned}$$

Define

$$w_{i+1}(x, t) = 5^{-(\sigma+\alpha)(i+1)} v_{i+1}(5^{i+1}x, 5^{(i+1)\sigma}t).$$

Then, (2.8), (2.9), (2.10) and (2.11) hold for $i+1$.

Moreover, for $(x, t) \in B_{4-2\tau} \times [-4\sigma + \tau^\sigma, 0]$, we have, similar to (2.21),

$$\begin{aligned}& |\partial_t W(x, t) - \partial_t v_{i+1}(x, t) - \int_{\mathbb{R}^n} \delta(W(x, y; t) - v_{i+1}(x, y; t)) K^{(i+1)}(x, y; t) dy| \\ & \leq |5^{\alpha(i+1)}(f(5^{-(i+1)}x, 5^{-(i+1)\sigma}t) - f(0, 0))| \\ & + \sum_{l=0}^{i+1} \int_{\mathbb{R}^n} 5^{(\sigma+\alpha)(i+1)} |\delta w_l(5^{-(i+1)}x, 5^{-(i+1)}y; 5^{-(i+1)\sigma}t)| \cdot \\ & |K^{i+1}(0, y; t) - K^{(i+1)}(x, y; t)| dy \\ & \leq \tau^{-\sigma} \eta_2 \leq 5^{-(\sigma+\alpha)} \tau^{-\sigma},\end{aligned}$$

where we used (2.23) in the second inequality. Thus, it follows from (2.18) and (2.24) that

$$[W - v_{i+1}]_{C^{\alpha_1}(B_{4-3\tau} \times [-4\sigma + 2\tau^\sigma, 0])} \leq 8c_1 5^{-(\sigma+\alpha)} \tau^{-4}.$$

Thus, (2.12) hold for $i+1$ as well. This finishes the proof of the claim. \square

A corollary of Theorem 1.1 would be the Schauder estimates for elliptic equations. If we consider the linear elliptic integro-differential equation

$$Lu(x) = f(x) \quad \text{in } B_5 \tag{2.25}$$

where

$$Lu(x) := \int_{\mathbb{R}^n} \delta u(x, y) K(x, y) dy, \tag{2.26}$$

$\delta u(x, y) = u(x+y) + u(x-y) - 2u(x)$, $K(x, y) \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$. We assume that

$$\int_{\mathbb{R}^n} |K(x, y) - K(0, y)| \min(|y|^2, r^2) dy \leq \Lambda |x|^\alpha r^{2-\sigma} \tag{2.27}$$

for all $r \in (0, 1]$, $x \in B_5$, and

$$|f(x) - f(0)| \leq M_f |x|^\alpha \quad \text{and} \quad |f(x)| \leq M_f \quad \forall x \in B_5 \tag{2.28}$$

for some positive constant M_f .

Corollary 2.5. *Let $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$ with $2 > \sigma \geq \sigma_0 > 0$. Let $\alpha \in (0, 1)$ such that $|\sigma + \alpha - j| \geq \varepsilon_0$ for some $\varepsilon_0 > 0$, where $j = 1, 2, 3$. Suppose that (2.27) and (2.28) hold. If u is a viscosity solution of (2.25), then there exists a polynomial $P(x)$ of degree $[\sigma + \alpha]$ such that for $x \in B_1$*

$$\begin{aligned} |u(x) - P(x)| &\leq C (\|u\|_{L^\infty(\mathbb{R}^n)} + M_f) |x|^{\sigma+\alpha}; \\ |\nabla^j P(0)| &\leq C (\|u\|_{L^\infty(\mathbb{R}^n)} + M_f), \quad j = 0, 1, [\sigma + \alpha], \end{aligned} \quad (2.29)$$

where C is a positive constant depending only on $\lambda, \Lambda, n, \sigma_0, \alpha$ and ε_0 .

The constant C in (2.29) does not blow up as $\sigma \rightarrow 2$, but it will blow up as $\sigma + \alpha$ approaches to integers.

2.3 Hölder estimates in space-time for spatial derivatives

Another corollary of the pointwise Schauder estimate (1.9) is the following uniform (in t) interior Schauder estimates in spatial variables.

We say that $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma, \alpha)$ if $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma)$ and

$$\int_{\mathbb{R}^n} |K(x_1, y, t_1) - K(x_2, y, t_2)| \min(|y|^2, r^2) dy \leq \Lambda(|x_1 - x_2|^\alpha + |t_1 - t_2|^{\frac{\alpha}{\sigma}}) r^{2-\sigma} \quad (2.30)$$

for all $r \in (0, 1]$, $x_1, x_2 \in B_5, t_1, t_2 \in (-5^\sigma, 0]$. We also assume that

$$|f(x_1, t_1) - f(x_2, t_2)| \leq M_f(|x_1 - x_2|^\alpha + |t_1 - t_2|^{\frac{\alpha}{\sigma}}), \quad |f(x_1, t_1)| \leq M_f \quad (2.31)$$

for all $x_1, x_2 \in B_5, t_1, t_2 \in (-5^\sigma, 0]$ and some positive constant M_f .

Corollary 2.6. *Let $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma, \alpha)$ with $2 > \sigma \geq \sigma_0 > 0$. Let $\alpha \in (0, 1)$ such that $|\sigma + \alpha - j| \geq \varepsilon_0$ for some $\varepsilon_0 > 0$, where $j = 1, 2, 3$. Suppose that (2.31) holds. If u is a viscosity solution of (1.1), then $u(\cdot, t) \in C_x^{\sigma+\alpha}(B_1)$ for all $t \in (-1, 0]$, and there exists a constant C depending only on $\lambda, \Lambda, n, \sigma_0, \alpha, \varepsilon_0$, such that*

$$\sup_{t \in [-1, 0]} \|u(\cdot, t)\|_{C^{\sigma+\alpha}(B_1)} \leq C (\|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + M_f). \quad (2.32)$$

Corollary 2.6 follows from Theorem 1.1 and standard translation arguments. Once we know the optimal regularity estimates of $\nabla_x u$ or $\nabla_x^2 u$ in the spatial variables, we can also obtain their regularity estimates in the time variable.

We say that $K \in \mathcal{L}_3(\lambda, \Lambda, \sigma, \alpha)$ if $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma, \alpha)$ and $|\nabla_y^3 K(x, y; t)| \leq \Lambda |y|^{-n-\sigma-3}$.

Corollary 2.7. *Let $K \in \mathcal{L}_2(\lambda, \Lambda, \sigma, \alpha)$ with $2 > \sigma \geq \sigma_0 > 0$. Let $\alpha \in (0, 1)$ such that $|\sigma + \alpha - j| \geq \varepsilon_0$ for some $\varepsilon_0 > 0$, where $j = 1, 2, 3$. Suppose that (2.31) holds. If u is a viscosity solution of (1.1), then*

$$|u(x_1, t_1) - u(x_2, t_2)| \leq C_1 (\|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + M_f) (|t_1 - t_2| + |x_1 - x_2|) \quad (2.33)$$

for all $x_1, x_2 \in B_1$, $t_1, t_2 \in (-1, 0]$; if $1 < \sigma + \alpha < 2$, there holds

$$\|\nabla_x u\|_{C_{x,t}^{\sigma+\alpha-1, \frac{\sigma+\alpha-1}{\sigma}}(B_1 \times [-1, 0])} \leq C_2 (\|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + M_f); \quad (2.34)$$

if $\sigma + \alpha > 2$ and $K \in \mathcal{L}_3(\lambda, \Lambda, \sigma, \alpha)$, there hold (2.34) and

$$\|\nabla_x^2 u\|_{C_{x,t}^{\sigma+\alpha-2, \frac{\sigma+\alpha-2}{\sigma}}(B_1 \times [-1, 0])} \leq C_3 (\|u\|_{L^\infty(\mathbb{R}^n \times (-5^\sigma, 0])} + M_f), \quad (2.35)$$

where C_1, C_2, C_3 are positive constants depending only on $\lambda, \Lambda, n, \sigma_0, \alpha, \varepsilon_0$.

In particular, the constants C, C_1, C_2, C_3 in (2.32)-(2.35) do not blow up as $\sigma \rightarrow 2^-$.

Lemma 2.8. *Let $v \in C(\mathbb{R}^n \times [-5^\sigma, 0])$ satisfies (2.13) in viscosity sense. Then v is locally Lipschitz in time. Moreover, for $t \in [-1, 0]$ there holds*

$$|v(0, t) - v(0, 0)| \leq C (\|v\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, 0])} + |f(0, 0)|) |t|.$$

By using (2.17) and the equation (2.13) itself, this estimate is clear if we consider it as a priori estimate.

Proof of Lemma 2.8. Let $t_0 \in (0, 4^\sigma)$. To proceed, we smooth v by using a mollifier $\eta_\varepsilon(x, t)$, and let $g_\varepsilon = \eta_\varepsilon * v$. Let v_ε be the solution of

$$\begin{aligned} \partial_t v_\varepsilon - L_0 v_\varepsilon &= f(0, 0) \quad \text{in } B_4 \times (-4^\sigma, -t_0] \\ v_\varepsilon &= g_\varepsilon \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\sigma, -t_0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}). \end{aligned}$$

It follows from Theorem 4.1 in [10] that $\partial_t v_\varepsilon$ is Hölder continuous in space-time. By Proposition 2.1 and Proposition 2.4, we know that v_ε is C^2 in x . Thus, v_ε satisfies its equation in the classical sense. By the equation of v_ε ,

$$\begin{aligned} &\|\partial_t v_\varepsilon\|_{L^\infty(B_1 \times [-1, -t_0])} \\ &\leq \|L_0 v_\varepsilon\|_{L^\infty(B_1 \times [-1, -t_0])} + |f(0, 0)| \\ &\leq C(\|\nabla_x^2 v_\varepsilon\|_{L^\infty(B_3 \times [-1, -t_0])} + \|v_\varepsilon\|_{L^\infty(\mathbb{R}^n \times [-1, -t_0])} + |f(0, 0)|) \\ &\leq C(\|v_\varepsilon\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, -t_0])} + |f(0, 0)|) \\ &\leq C(\|v\|_{L^\infty(\mathbb{R}^n \times [-4^\sigma, -t_0])} + |f(0, 0)|), \end{aligned}$$

where the estimates in Proposition 2.1 and Proposition 2.4 are used in the third inequality. Meanwhile, by the Hölder interior estimates, we have that v_ε locally uniformly converges to some continuous function w . By the stability result Theorem 5.3 in [10], w is a viscosity solution of

$$\begin{aligned} \partial_t w - L_0 w &= f(0, 0) \quad \text{in } B_4 \times (-4^\sigma, -t_0] \\ w &= v \quad \text{in } ((\mathbb{R}^n \setminus B_4) \times [-4^\sigma, -t_0]) \cup (\mathbb{R}^n \times \{t = -4^\sigma\}). \end{aligned}$$

Hence $w \equiv v$, and thus, we have that

$$\begin{aligned} |v(0, t) - v(0, -t_0)| &= \lim_{\varepsilon \rightarrow 0} |v_\varepsilon(0, t) - v_\varepsilon(0, -t_0)| \\ &\leq C \left(\|v\|_{L^\infty(\mathbb{R}^n \times [-4\sigma, 0])} + |f(0, 0)| \right) |t + t_0|. \end{aligned}$$

We finish the proof by sending $t_0 \rightarrow 0$. \square

Remark 2.9. *Indeed, by similar arguments and the integration by parts technique used in the proof of Proposition 2.1 one can also show that $\nabla_x v$ is Lipschitz in time, as well as $\nabla_x^2 v$ if $K \in \mathcal{L}_3(\lambda, \Lambda, \sigma, \alpha)$ and $\sigma > 1$ (so that we have estimates for $\nabla_x^4 v$). We omit the proof here.*

Proof of Corollary 2.7. If we let w_l be as in the proof of Theorem 1.1, then by Lemma 2.8 and Remark 2.9, we have that $w_l, \nabla_x w_l$ is Lipschitz in time, as well as $\nabla_x^2 w_l$ provided that $K \in \mathcal{L}_3(\lambda, \Lambda, \sigma, \alpha)$. By Corollary 2.6, we may assume that $x_1 = x_2 = 0, t_2 = 0$. Suppose that $-5^{i\sigma} \leq t < -5^{(i+1)\sigma}$. Then we have

$$\begin{aligned} &|u(0, t) - u(0, 0)| \\ &\leq |u(0, t) - \sum_{l=0}^i w_l(0, t)| + |u(0, 0) - \sum_{l=0}^i w_l(0, 0)| + \sum_{l=0}^i |w_l(0, t) - w_l(0, 0)| \\ &= 2 \cdot 5^{-(\sigma+\alpha)(i+1)} + C \sum_{l=0}^i 5^{-\alpha l} |t| \leq C_1 |t|, \end{aligned}$$

which proves (2.33).

Suppose that $1 < \sigma + \alpha < 2$. We have

$$\begin{aligned} |\nabla_x u(0, t) - \nabla_x u(0, 0)| &\leq |\nabla_x u(0, t) - \sum_{l=0}^i \nabla_x w_l(0, t)| \\ &\quad + \sum_{l=0}^i |\nabla_x w_l(0, t) - \nabla_x w_l(0, 0)| \\ &\quad + |\nabla_x u(0, 0) - \sum_{l=0}^i \nabla_x w_l(0, 0)| := I_1 + I_2 + I_3. \end{aligned}$$

Since $\nabla_x u(0, 0) = \sum_{l=0}^\infty \nabla_x w_l(0, 0)$, we have

$$|I_3| \leq \sum_{l=i+1}^\infty c_2 5^{-(\sigma+\alpha-1)l} \leq c_2 |t|^{\frac{\sigma+\alpha-1}{\sigma}} \frac{1}{1 - 5^{-(\sigma+\alpha-1)}}.$$

By the equation of w_l , Lemma 2.8, Remark 2.9 and (2.11), we have

$$|\nabla_x w_l(0, t) - \nabla_x w_l(0, 0)| \leq C 5^{(1-\alpha)l} |t|.$$

Then

$$|I_2| \leq C|t| \frac{5^{(1-\alpha)(i+1)}}{5^{1-\alpha} - 1} \leq C|t|^{\frac{\sigma+\alpha-1}{\sigma}} \frac{5^{1-\alpha}}{5^{1-\alpha} - 1}.$$

Meanwhile, it follows from the estimate (1.9) that

$$|u(x, t) - u(0, t) - \nabla_x u(0, t)x| \leq C|x|^{\sigma+\alpha}.$$

For $5^{-(i+1)} \leq |x| < 5^{-i}$, we have, by triangle inequality,

$$\begin{aligned} & |\nabla_x u(0, t)x - \sum_{l=0}^i \nabla_x w_l(0, t)x| \\ & \leq |u(x, t) - \sum_{l=0}^i w_l(x, t)| + |\sum_{l=0}^i (w_l(x, t) - w_l(0, t) - \nabla_x w_l(0, t)x)| \\ & \quad + |u(0, t) + \nabla_x u(0, t)x - u(x, t)| + |\sum_{l=0}^i w_l(0, t) - u(0, t)| \\ & \leq 5^{-(\sigma+\alpha)(i+1)} + c_2 \sum_{l=0}^i 5^{-(\sigma+\alpha-2)l} |x|^2 + C|x|^{\sigma+\alpha} + 5^{-(\sigma+\alpha)(i+1)} \end{aligned}$$

Thus

$$\begin{aligned} |I_1| &= |\nabla_x u(0, t) - \sum_{l=0}^i \nabla_x w_l(0, t)| \leq (C + \frac{1}{5^{2-\sigma-\alpha} - 1}) 5^{-(\sigma+\alpha-1)(i+1)} \\ &\leq (C + \frac{1}{5^{2-\sigma-\alpha} - 1}) |t|^{\frac{\sigma+\alpha-1}{\sigma}}. \end{aligned}$$

Hence, we have shown that

$$|\nabla_x u(0, t) - \nabla_x u(0, 0)| \leq C_2 |t|^{\frac{\sigma+\alpha-1}{\sigma}}.$$

Suppose that $2 < \sigma + \alpha < 3$. We have

$$\begin{aligned} |\nabla_x^2 u(0, t) - \nabla_x^2 u(0, 0)| &\leq |\nabla_x^2 u(0, t) - \sum_{l=0}^i \nabla_x^2 w_l(0, t)| \\ &\quad + \sum_{l=0}^i |\nabla_x^2 w_l(0, t) - \nabla_x^2 w_l(0, 0)| \\ &\quad + |\nabla_x^2 u(0, 0) - \sum_{l=0}^i \nabla_x^2 w_l(0, 0)| := II_1 + II_2 + II_3. \end{aligned}$$

Since $\nabla_x^2 u(0, 0) = \sum_{l=0}^{\infty} \nabla_x^2 w_l(0, 0)$, we have

$$|II_3| \leq c_2 |t|^{\frac{\sigma+\alpha-2}{\sigma}} \frac{1}{1 - 5^{-(\sigma+\alpha-2)}}.$$

Meanwhile, it follows from the estimate (1.9) that

$$|u(x, t) - u(0, t) - \nabla_x u(0, t)x - \frac{1}{2}x^T \nabla_x^2 u(0, t)x| \leq C|x|^{\sigma+\alpha}.$$

By triangle inequality and the estimate for I_1 , we have, for $5^{-(i+1)} \leq |x| < 5^{-i}$

$$\begin{aligned} & \frac{1}{2} |x^T \nabla_x^2 u(0, t)x - x^T \sum_{l=0}^i \nabla_x^2 w_l(0, t)x| \\ & \leq |u(x, t) - \sum_{l=0}^i w_l(x, t)| + |u(0, t) + \nabla_x u(0, t)x + \frac{1}{2}x^T \nabla_x^2 u(0, t)x - u(x, t)| \\ & \quad + |\sum_{l=0}^i (w_l(x, t) - w_l(0, t) - \nabla_x w_l(0, t)x - \frac{1}{2}x^T \nabla_x^2 w_l(0, t)x)| \\ & \quad + |\sum_{l=0}^i w_l(0, t) - u(0, t)| + |\sum_{l=0}^i \nabla_x w_l(0, t)x - \nabla_x u(0, t)x| \\ & \leq 5^{-(\sigma+\alpha)(i+1)} + c_2 \sum_{l=0}^i 5^{-(\sigma+\alpha-3)l} |x|^3 + C|x|^{\sigma+\alpha} + 5^{-(\sigma+\alpha)(i+1)} \\ & \leq 2 \cdot 5^{-(\sigma+\alpha)(i+1)} + c_2 \frac{5^{-(\sigma+\alpha)(i+1)}}{5^{2-\sigma-\alpha} - 1} + C5^{-(\sigma+\alpha)i} + (C + \frac{1}{5^{2-\sigma-\alpha} - 1})5^{-(\sigma+\alpha)(i+1)}. \end{aligned}$$

Thus

$$|II_1| = |\nabla_x u(0, t) - \sum_{l=0}^i \nabla_x w_l(0, t)| \leq C5^{-(\sigma+\alpha-2)(i+1)} \leq C|t|^{\frac{\sigma+\alpha-2}{\sigma}}.$$

By the equation of w_l , Lemma 2.8, Remark 2.9 and (2.11), we have

$$|\nabla_x^2 w_l(0, t) - \nabla_x^2 w_l(0, 0)| \leq C_3 5^{(2-\alpha)l} |t|$$

provided that $K \in \mathcal{L}_3(\lambda, \Lambda, \sigma, \alpha)$. Then

$$|I_2| \leq C|t| \frac{5^{(2-\alpha)(i+1)}}{5^{2-\alpha} - 1} \leq C|t|^{\frac{\sigma+\alpha-2}{\sigma}} \frac{5^{2-\alpha}}{5^{2-\alpha} - 1}.$$

Thus, by combining the estimates for II_1, II_2, II_3 , we have that

$$|\nabla_x^2 u(0, t) - \nabla_x^2 u(0, 0)| \leq C|t|^{\frac{\sigma+\alpha-2}{\sigma}}.$$

This completes the proof of Corollary 2.7. □

If we do not assume $K \in \mathcal{L}_3(\lambda, \Lambda, \sigma, \alpha)$ when $\sigma + \alpha > 2$, we have that $\nabla_x^2 u$ is of C^β in the time variable for some $\beta > 0$. This is because $\nabla_x^2 u$ is Hölder continuous in x and $\nabla_x u$ is Hölder continuous in t , which implies that $\nabla_x^2 u$ is Hölder continuous in t as well; see Lemma 3.1 on page 78 in [18].

A Approximation lemmas

Our proof of Schauder estimates uses perturbative arguments, and we need the following two approximation lemmas, which are variants of Theorem 5.6 in [10] (Lemma 7 in [6] in elliptic cases). We will do a few modifications for our own purposes, and we decide to include them in this appendix for completeness and convenience. If it is just for our particular linear equations, those approximation lemmas can be simplified much. But we would like to include nonlinear equations as well in this step.

To start with, we recall some definitions and notations about nonlocal operators, which can be found in [9, 10] for parabolic cases and in [5, 6] for elliptic cases. Let $\sigma_0 \in (0, 2)$ be fixed, and $\omega(y) = (1 + |y|^{n+\sigma_0})^{-1}$. We say $u \in L^1(\mathbb{R}^n, \omega)$ if $\int_{\mathbb{R}^n} |u(y)|\omega(y)dy < \infty$. We say that $u \in C(a, b; L^1(\omega))$ if $u(\cdot, t) \in L^1(\mathbb{R}^n, \omega)$ for every $t \in (a, b)$, and $\|u(\cdot, t_1) - u(\cdot, t_2)\|_{L^1(\mathbb{R}^n, \omega)} \rightarrow 0$ as $t_1 \rightarrow t_2$ for every $t_2 \in (a, b)$, and we denote $\|u\|_{C(a, b; L^1(\omega))} = \sup_{t \in (a, b)} \|u(\cdot, t)\|_{L^1(\mathbb{R}^n, \omega)}$.

Nonlocal (continuous) operators I are defined as “black boxes” in Definitions 3.3 and 3.6 in [9] such that, rough speaking, if u is a test function at (x, t) , the Iu is continuous near (x, t) . In our case, they are just linear operators of the form (1.2) with some continuity assumptions on K in (x, t) . Sometimes, we also write $Iu(x, t)$ as $I(u, x, t)$ for convenience especially when dealing with $I(u + v)$.

An operator is translation invariant if $\tau_{(z, s)} Iu = I(\tau_{(z, s)} u)$ where $\tau_{(z, s)}$ is the translation operator $\tau_{(z, s)} u(x, t) = u(x - z, t - s)$.

Given such a nonlocal operator I defined on $\Omega \times (-T, 0]$, a norm $\|I\|$ was defined in Definition 5.3 in [10]. Here, we also define a (weaker) norm $\|I\|_*$ for our own purpose,

$$\begin{aligned} \|I(t)\|_* &:= \sup\{|Iu(x, t)|/(1 + M) : x \in \Omega, \|u(\cdot, t)\|_{\mathbb{R}^n} \leq M \\ &\quad |u(x + y, t) - u(x, t) - y \cdot \nabla_x u(x, t)| \leq M|y|^2 \text{ for every } y \in B_1\}, \end{aligned} \quad (\text{A.1})$$

and $\|I\|_* = \sup_{t \in (-T, 0]} \|I(t)\|_*$.

We say that a nonlocal operator I is uniformly elliptic with respect to $\mathcal{L}_0(\lambda, \Lambda, \sigma)$, which will be written as $\mathcal{L}_0(\sigma)$ for short, if

$$\mathcal{M}_{\mathcal{L}_0(\sigma)}^- v(x, t) \leq I(u + v, x, t) - I(u, x, t) \leq \mathcal{M}_{\mathcal{L}_0(\sigma)}^+ v(x, t), \quad (\text{A.2})$$

where

$$\mathcal{M}_{\mathcal{L}_0(\sigma)}^- v(x, t) = \inf_{L \in L_0(\sigma)} Lv(x, t) = (2 - \sigma) \int_{\mathbb{R}^n} \frac{\lambda \delta v(x, y; t)^+ - \Lambda \delta v(x, y; t)^-}{|y|^{n+\sigma}} dy$$

$$\mathcal{M}_{\mathcal{L}_0(\sigma)}^+ v(x, t) = \sup_{L \in L_0(\sigma)} Lv(x, t) = (2 - \sigma) \int_{\mathbb{R}^n} \frac{\Lambda \delta v(x, y; t)^+ - \lambda \delta v(x, y; t)^-}{|y|^{n+\sigma}} dy.$$

It is also convenient to define the limit operators when $\sigma \rightarrow 2$ as

$$\begin{aligned} \mathcal{M}_{\mathcal{L}_0(2)}^- v(x, t) &= \lim_{\sigma \rightarrow 2} \mathcal{M}_{\mathcal{L}_0(\sigma)}^- v(x, t) \\ \mathcal{M}_{\mathcal{L}_0(2)}^+ v(x, t) &= \lim_{\sigma \rightarrow 2} \mathcal{M}_{\mathcal{L}_0(\sigma)}^+ v(x, t). \end{aligned}$$

It has been explained in [6] that $\mathcal{M}_{\mathcal{L}_0(2)}^+$ is a second order uniformly elliptic operator, whose ellipticity constants $\tilde{\lambda}$ and $\tilde{\Lambda}$ depend only λ, Λ and the dimension n . Moreover, $\mathcal{M}_{\mathcal{L}_0(2)}^+ v \leq \mathcal{M}^+(\nabla^2 v)$, where $\mathcal{M}^+(\nabla^2 v)$ is the second order Pucci operator with ellipticity constants $\tilde{\lambda}$ and $\tilde{\Lambda}$. Similarly, we also have corresponding relations for $\mathcal{M}_{\mathcal{L}_0(2)}^-$.

For compactness arguments, we shall use the concept of the weak convergence of nonlocal operators, which can be found in Definition 5.1 in [10] (Definition 41 in [6] in the elliptic cases).

Lemma A.1. *For some $\sigma \geq \sigma_0 > 0$ we consider nonlocal continuous operators I_0, I_1 and I_2 uniformly elliptic with respect to $\mathcal{L}_0(\sigma)$. Assume also that I_0 is translation invariant and $I_0 0 = 1$.*

Given $M_1 > 0$, a modulus of continuity ρ and $\varepsilon > 0$, there exist η_1 (small, independent of σ) and R (large, independent of σ) so that if u, v, I_0, I_1 and I_2 satisfy

$$\begin{aligned} v_t - I_0(v, x, t) &= 0 \quad \text{in } B_1 \times (-1, 0], \\ u_t - I_1(u, x, t) &\leq \eta_1 \quad \text{in } B_1 \times (-1, 0], \\ u_t - I_2(u, x, t) &\geq -\eta_1 \quad \text{in } B_1 \times (-1, 0], \end{aligned}$$

in viscosity sense, and

$$\begin{aligned} \|I_1 - I_0\|_* &\leq \eta_1 \quad \text{in } B_1 \times (-1, 0], \\ \|I_2 - I_0\|_* &\leq \eta_1 \quad \text{in } B_1 \times (-1, 0], \\ u &= v \quad \text{in } ((\mathbb{R}^n \setminus B_1) \times [-1, 0]) \cup (B_1 \times \{t = -1\}), \\ |u| &\leq M_1 \quad \text{in } \mathbb{R}^n \times [-1, 0], \end{aligned}$$

and for every $(x, t) \in ((B_R \setminus B_1) \times (-1, 0]) \cup (B_R \times \{t = -1\})$ and $(y, s) \in ((\mathbb{R}^n \setminus B_1) \times (-1, 0]) \cup (\mathbb{R}^n \times \{t = -1\})$,

$$|u(x, t) - u(y, s)| \leq \rho(|x - y| \vee |t - s|),$$

then $|u - v| \leq \varepsilon$ in $B_1 \times (-1, 0]$.

Proof. It follows from the proof of Theorem 5.6 in [10] with modifications. But since the choice of norms are different, we include the proof for completeness. We argue by contradiction. Suppose the above lemma was false. Then there would be sequences $\sigma_k, I_0^{(k)}, I_1^{(k)}, I_2^{(k)}, \eta_k, u_k, v_k$ such that

$\sigma_k \rightarrow \sigma \in [\sigma_0, 2]$, $\eta_k \rightarrow 0$ and all the assumptions of the lemma are valid, but $\sup_{B_1 \times (-1, 0]} |u_k - v_k| \geq \varepsilon$.

Since $I_0^{(k)}$ is a sequence of uniformly elliptic translation invariant operators with respect to $\mathcal{L}(\sigma_k)$, by Theorem 5.5 in [10] (and its proof) that we can take a subsequence, which is still denoted as $I_0^{(k)}$, that converges weakly to some nonlocal operator I_0 , and I_0 is also translation invariant uniformly elliptic with respect to the class $\mathcal{L}_0(\sigma)$.

It follows from the boundary regularity Theorem 3.2 in [10] that u_k and v_k have a modulus of continuity, uniform in k , in $\overline{B}_1 \times [-1, 0]$. Thus, u_k and v_k have a uniform (in k) modulus of continuity on $B_{R_k} \times [-1, 0]$ with $R_k \rightarrow \infty$. We have subsequences of $\{u_k\}$ and $\{v_k\}$, which will be still denoted as $\{u_k\}$ and $\{v_k\}$, converge locally uniformly in $\mathbb{R}^n \times [-1, 0]$ to u and v , as well as in $C(-1, 0, L^1(\mathbb{R}^n, \omega))$ by dominated convergence theorem, respectively. Moreover, $u = v$ in $((\mathbb{R}^n \setminus B_1) \times [-1, 0]) \cup (B_1 \times \{t = -1\})$, and $\sup_{B_1 \times (-1, 0]} |u - v| \geq \varepsilon$.

In the following, we are going to show in viscosity sense that

$$u_t - I_0(u, x, t) = 0 = v_t - I_0(v, x, t) \quad \text{in } B_1 \times (-1, 0]. \quad (\text{A.3})$$

Since I_0 translation invariant and $u = v$ in $((\mathbb{R}^n \setminus B_1) \times [-1, 0]) \cup (B_1 \times \{t = -1\})$, we can conclude from Corollary 3.1 in [10] that $u \equiv v$ in $B_1 \times [-1, 0]$, which is a contradiction.

The second equality of (A.3) follows from Theorem 5.3 in [10]. To prove the first equality of (A.3), let p be a second order parabolic polynomial touching u from below at a point $(x, t) \in B_1 \times (-1, 0]$ in a neighborhood $B_r(x) \times (t - r, t]$. Since u_k converges uniformly to u in $\overline{B}_1 \times [-1, 0]$, for large k , we can find $(x_k, t_k) \in B_r(x) \times (t - r, t]$ and d_k so that $p + d_k$ touch u_k at (x_k, t_k) . Furthermore, $(x_k, t_k) \rightarrow (x, t)$ and $d_k \rightarrow 0$ as $k \rightarrow \infty$. Since $\partial_t u_k - I_2^{(k)}(u_k, x) \geq -\eta_k$, if we let

$$w_k(y, s) = \begin{cases} p + d_k & \text{in } B_r(x) \times (t - r, t]; \\ u_k & \text{in } (\mathbb{R}^n \setminus B_r(x)) \times (B_r(x) \times \{s = t - r\}), \end{cases}$$

we have $\partial_t w_k(x_k, t_k) - I_2^{(k)}(w_k, x_k, t_k) \geq -\eta_k$, and

$$w = \lim_{k \rightarrow \infty} w_k = \begin{cases} p & \text{in } B_r(x) \times (t - r, t]; \\ u & \text{in } (\mathbb{R}^n \setminus B_r(x)) \times (B_r(x) \times \{s = t - r\}). \end{cases}$$

Let $(z, s) \in B_{r/4}(x) \times (t - r/4, t]$. We have

$$\begin{aligned} & |I_2^{(k)}(w_k, z, s) - I_0(w, z, s)| \\ & \leq |I_2^{(k)}(w_k, z, s) - I_2^{(k)}(w, z, s)| + |I_2^{(k)}(w, z, s) - I_0(w, z, s)| \\ & \leq \sup_{L \in \mathcal{L}(\sigma_k)} |L(w_k - w)(z, s)| + |I_2^{(k)}(w, z, s) - I_0(w, z, s)| \\ & \leq \int_{\mathbb{R}^n \setminus B_{r/2}} \frac{2\Lambda |\delta(w_k - w)(z, y, s)|}{|y|^{n+\sigma_k}} dy + |I_2^{(k)}(w, z, s) - I_0^{(k)}(w, z, s)| \\ & \quad + |I_0^{(k)}(w, z, s) - I_0(w, z, s)|. \end{aligned}$$

Since u_k are uniformly bounded in $\mathbb{R}^n \times [0, 1]$, by dominated convergence theorem, the first term goes to 0 as $k \rightarrow \infty$. Moreover, the convergence is uniform in (z, s) . Meanwhile, since $\|I_2^{(k)} - I_0^{(k)}\|_* \rightarrow 0$ in $B_1 \times (-1, 0]$ and w is bounded, we have that the second goes to 0 uniformly for $(z, s) \in B_{r/4}(x) \times (t - r/4, t]$. Since $I_0^{(k)}$ converges weakly to I_0 , the third term also goes to zero uniformly for $(z, s) \in B_{r/4}(x) \times (t - r/4, t]$. Therefore, $I_2^{(k)}(w_k, z, s) \rightarrow I_0(w, z, s)$ uniformly in $(z, s) \in B_{r/4}(x) \times (t - r/4, t]$. Since $I_0 w$ is continuous in $B_r(x)$, we can compute that

$$\begin{aligned} & |I_2^{(k)}(w_k, x_k, t_k) - I_0(w, x, t)| \\ & \leq |I_2^{(k)}(w_k, x_k, t_k) - I_0(w, x_k, t_k)| + |I_0(w, x_k, t_k) - I_0(w, x, t)| \rightarrow 0 \end{aligned}$$

as $k \rightarrow \infty$. Since $\partial_t w_k(x_k, t_k) - I_2^{(k)}(w_k, x_k, t_k) \geq -\eta_k$ and $\partial_t w_k(x_k, t_k) \rightarrow \partial_t w(x, t)$, it follows that $w_t(x, t) - I_0(w, x, t) \geq 0$. Thus, $u_t(x, t) - I_0(u, x, t) \geq 0$ in viscosity sense. Similarly, we can show that $u_t(x, t) - I_0(u, x, t) \leq 0$ in viscosity sense. This finishes the proof of the first equality of (A.3). \square

Lemma A.2. *For some $\sigma \geq \sigma_0 > 0$ we consider nonlocal continuous operators I_0 , I_1 and I_2 uniformly elliptic with respect to $\mathcal{L}_0(\sigma)$. Assume also that I_0 is translation invariant and $I_0 0 = 0$.*

Given $M_2, M_3 > 0$, $\beta \in (0, 1)$, and $\varepsilon > 0$, there exists η_2 (small) so that if u, v, I_0, I_1 and I_2 satisfy

$$\begin{aligned} v_t - I_0(v, x, t) &= 0 \quad \text{in } B_1 \times (-1, 0], \\ u_t - I_1(u, x, t) &\leq \eta_2 \quad \text{in } B_1 \times (-1, 0], \\ u_t - I_2(u, x, t) &\geq -\eta_2 \quad \text{in } B_1 \times (-1, 0], \end{aligned}$$

in viscosity sense, and

$$\begin{aligned} \|I_1 - I_0\|_* &\leq \eta_2 \quad \text{in } B_1 \times (-1, 0], \\ \|I_2 - I_0\|_* &\leq \eta_2 \quad \text{in } B_1 \times (-1, 0], \\ u &= v \quad \text{in } ((\mathbb{R}^n \setminus B_1) \times [-1, 0]) \cup (B_1 \times \{t = -1\}), \\ |u| &\leq M_2 \quad \text{in } \mathbb{R}^n \times [-1, 0], \\ u &\equiv 0 \quad \text{in } (\mathbb{R}^n \setminus B_2) \times [-1, 0], \\ [u]_{C^\beta(B_{2-\tau} \times [-1, 0])} &\leq M_3 \tau^{-4} \text{ for all } \tau \in (0, 1), \end{aligned}$$

then $|u - v| \leq \varepsilon$ in B_1 .

Proof. This lemma can be proved similarly to Lemma A.1. Suppose the above lemma was false. Then there would be sequences $\sigma_k, I_0^{(k)}, I_1^{(k)}, I_2^{(k)}, \eta_k, u_k, v_k$ such that $\sigma_k \rightarrow \sigma \in [\sigma_0, 2]$, $\eta_k \rightarrow 0$ and all the assumptions of the lemma are valid, but $\sup_{B_1} |u_k - v_k| \geq \varepsilon$.

Since $I_0^{(k)}$ is a sequence of uniformly elliptic operators, we can take a subsequence, which is still denoted as $I_0^{(k)}$, that converges weakly to some nonlocal operator I_0 , and I_0 is also translation invariant and elliptic with respect to the class $\mathcal{L}_0(\sigma)$.

By our assumptions, it is clear that, up to a subsequence, u_k converges locally uniformly in $B_2 \times [-1, 0]$. Since $u_k \equiv 0$ in $(\mathbb{R}^n \setminus B_2) \times [-1, 0]$, it converges almost everywhere to some function u in $\mathbb{R}^n \times [-1, 0]$. By dominated convergence theorem, u_k converges to u in $C(-1, 0; L^1(\mathbb{R}^n, \omega))$. Since v_k is bounded and has a fixed modulus continuity on $((B_{3/2} \setminus B_1) \times [-1, 0]) \cup B_1 \times \{t = -1\}$, then by Theorem 3.2 in [10], there is another modulus continuity that extends to $\overline{B_1} \times [-1, 0]$. Hence, v_k converges uniformly in $\overline{B_{3/2}} \times [-1, 0]$, and thus, converges to some function v almost everywhere in $\mathbb{R}^n \times [-1, 0]$. Moreover, $u = v$ in $((\mathbb{R}^n \setminus B_1) \times [-1, 0]) \cup (B_1 \times \{t = -1\})$, and $\sup_{B_1 \times (0, 1)} |u - v| \geq \varepsilon$.

It follows from the proof of (A.3) that u and v solve the same equation $u_t - I_0(u, x, t) = 0 = v_t - I_0(v, x, t)$ in $B_1 \times (-1, 0]$. Then $u = v$, which is a contradiction. \square

References

- [1] B. Barrera, A. Figalli and E. Valdinoci, *Bootstrap regularity for integro-differential operators, and its application to nonlocal minimal surfaces*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., to appear.
- [2] R.F. Bass, *Regularity results for stable-like operators*, J. Funct. Anal., **257** (2009), 2693–2722.
- [3] A. Brandt, *Interior Schauder estimates for parabolic differential- (or difference-) equations via the maximum principle*, Israel J. Math., **7** (1969), 254–262.
- [4] L.A. Caffarelli, *Interior a priori estimates for solutions of fully nonlinear equations*, Ann. of Math. (2), **130** (1989), 189–213.
- [5] L.A. Caffarelli and L. Silvestre, *Regularity theory for fully nonlinear integro-differential equations*, Comm. Pure Appl. Math., **62** (2009), 597–638.
- [6] L.A. Caffarelli and L. Silvestre, *Regularity results for nonlocal equations by approximation*, Arch. Ration. Mech. Anal., **200** (2011), 59–88.
- [7] L.A. Caffarelli and L. Silvestre, *The Evans-Krylov theorem for non local fully non linear equations*, Ann. of Math. (2), **174** (2011), 1163–1187.
- [8] L.A. Caffarelli, C. Chan and A. Vasseur, *Regularity theory for parabolic nonlinear integral operators*, J. Amer. Math. Soc., **24** (2011), 849–869.
- [9] H.A. Chang Lara and G. Davila, *Regularity for solutions of non local parabolic equations*, Calc. Var. Partial Differential Equations, **49** (2014), 139–172.
- [10] H.A. Chang Lara and G. Davila, *Regularity for solutions of non local parabolic equations II*, J. Differential Equations, **256** (2014), 130–156.
- [11] H. Dong and D. Kim, *Schauder estimates for a class of non-local elliptic equations*, Discrete Contin. Dyn. Syst., **33** (2013), 2319–2347.

- [12] H. Dong and S. Kim, *Partial Schauder estimates for second-order elliptic and parabolic equations*, Calc. Var. Partial Differential Equations, **40** (2011), 481–500.
- [13] M. Felsinger and M. Kassmann, *Local regularity for parabolic nonlocal operators*, Comm. Partial Differential Equations, **38** (2013), 1539–1573.
- [14] M. Kassmann and R. W. Schwab, *Regularity results for nonlocal parabolic equations*, arXiv:1305.5418.
- [15] B.F. Knerr, *Parabolic interior Schauder estimates by the maximum principle*, Arch. Ration. Mech. Anal., **75** (1980), 51–58.
- [16] D. Kriventsov, *$C^{1,\alpha}$ interior regularity for nonlinear nonlocal elliptic equations with rough kernels*, Comm. Partial Differential Equations, **38** (2013), 2081–2106.
- [17] N. V. Krylov and E. Priola, *Elliptic and parabolic second-order PDEs with growing coefficients*, Comm. Partial Differential Equations, **35** (2010), 1–22.
- [18] O. A. Ladyzenskaja, V. A. Solonnikov and N. N. Ural’ceva, “Linear and quasilinear equations of parabolic type”, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I. 1968.
- [19] Y.Y. Li and L. Nirenberg, *Estimates for elliptic system from composition material*, Comm. Pure Appl. Math., **56** (2003), 892–925.
- [20] G.M. Lieberman, *Intermediate Schauder theory for second order parabolic equations. IV. Time irregularity and regularity*, Differential Integral Equations, **5** (1992), 1219–1236.
- [21] L. Lorenzi, *Optimal Schauder estimates for parabolic problems with data measurable with respect to time*, SIAM J. Math. Anal., **32** (2000), 588–615.
- [22] R. Mikulevicius and H. Pragarauskas, *On the Cauchy problem for integro-differential operators in Hölder classes and the uniqueness of the martingale problem*, to appear in Potential Anal., [DOI:10.1007/s11118-013-9359-4].
- [23] J. Serra, *Regularity for fully nonlinear nonlocal parabolic equations with rough kernels*, arXiv:1401.4521.
- [24] E.M. Stein, “Singular integrals and differentiability properties of functions”. Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970.
- [25] G. Tian, and X.-J. Wang, *A priori estimates for fully nonlinear parabolic equations*, Int. Math. Res. Not. IMRN, 2013, 3857–3877.

Tianling Jin

Department of Mathematics, The University of Chicago, 5734 S. University Ave, Chicago, IL 60637, USA

Email: tj@math.uchicago.edu

Jingang Xiong

Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China

Email: jxiong@math.pku.edu.cn